

# Combinatorial theorems relative to a random set

David Conlon\*

## Abstract

We describe recent advances in the study of random analogues of combinatorial theorems.

## 1 Introduction

Random graphs have played an integral role in extremal combinatorics since they were first used by Erdős [30] to prove an exponential lower bound for Ramsey numbers. The binomial random graph  $G_{n,p}$  is a graph on  $n$  vertices where each of the  $\binom{n}{2}$  possible edges is chosen independently with probability  $p$ . In modern terminology, Erdős' result says that with high probability  $G_{n,1/2}$  contains no clique or independent set of order  $2\log_2 n$ . This then translates to a lower bound of  $2^{t/2}$  for the Ramsey number  $R(t)$  (this will be defined in Section 2).

Although there were several applications of random graphs prior to their work, the first systematic study of random graphs was undertaken by Erdős and Rényi [32, 33]. The concept of random graphs was also introduced independently by several other authors but, as explained by Bollobás [10], ‘the other authors were all concerned with enumeration problems and their techniques were essentially deterministic.’ Though it has its origins in applications to extremal combinatorics, the theory of random graphs is now a rich and self-sustaining area of study (see, for example, [10, 67]).

Suppose that  $\mathcal{P}$  is a graph property, that is, a family of graphs closed under isomorphism. In studying random graphs, we are usually concerned with determining the probability that  $G_{n,p}$  is in  $\mathcal{P}$  for some property  $\mathcal{P}$ . For many properties, this probability exhibits a phase transition as  $p$  increases, changing abruptly from 0 to 1. The crossover point is known as the threshold. Formally, we say that  $p^* := p^*(n)$  is a threshold for  $\mathcal{P}$  if

$$\lim_{n \rightarrow \infty} \mathbb{P}[G_{n,p} \text{ is in } \mathcal{P}] = \begin{cases} 0 & \text{if } p = o(p^*), \\ 1 & \text{if } p = \omega(p^*). \end{cases}$$

We note that, depending on the property  $\mathcal{P}$ , the probability could also collapse from 1 to 0 as  $p$  increases. However, for most properties considered in this paper, the behaviour is as above. To give some simple examples, the properties of being connected and having a Hamiltonian cycle are both known to have a threshold at  $p^* = \frac{\log n}{n}$ , while the property of containing a particular graph

---

\*Mathematical Institute, Oxford OX2 6GG, United Kingdom. Email: david.conlon@maths.ox.ac.uk. Research supported by a Royal Society University Research Fellowship.

$H$  has a threshold at  $n^{-1/m(H)}$ , where

$$m(H) = \max \left\{ \frac{e(H')}{v(H')} : H' \subseteq H \right\}.$$

This function reflects the fact that a graph appears once its densest subgraph does.

Since the late eighties, there has been a great deal of interest in determining thresholds for analogues of combinatorial theorems to hold in random graphs and random subsets of other sets such as the integers. To give an example, we say that a graph  $G$  is  $K_3$ -Ramsey if any 2-colouring of the edges of  $G$  contains a monochromatic triangle. One of the foundational results in this area, proved by Frankl and Rödl [40] and Łuczak, Ruciński and Voigt [86], then states that there exists  $C > 0$  such that if  $p > C/\sqrt{n}$  then

$$\lim_{n \rightarrow \infty} \mathbb{P}[G_{n,p} \text{ is } K_3\text{-Ramsey}] = 1.$$

Frankl and Rödl used this theorem to prove that there are  $K_4$ -free graphs which are  $K_3$ -Ramsey, a result originally due to Folkman [38]. However, this new method allowed one to prove reasonable bounds for the size of such graphs, something which was not possible with previous methods.

From this beginning, a large number of papers were written on sparse random analogues of combinatorial theorems. These included papers on analogues of Ramsey's theorem, Turán's theorem and Szemerédi's theorem, though in many cases these efforts met with only partial success. This situation has changed dramatically in recent years and there are now three distinct, general methods for proving sparse random analogues of combinatorial theorems, furnishing solutions for many of the outstanding problems in the area.

The first two of these methods were developed by Gowers and the author [24] and, independently, by Schacht [112] and Friedgut, Rödl and Schacht [45]. The third method was found later by Balogh, Morris and Samotij [6] and, independently, by Saxton and Thomason [111]. Broadly speaking, the method employed by Gowers and the author builds on the transference principle developed by Green and Tao [60] in their proof that the primes contain arbitrarily long arithmetic progressions; the method of Schacht and Friedgut, Rödl and Schacht extends a multi-round exposure technique used by Rödl and Ruciński [98] in their study of Ramsey's theorem in random graphs; and the third method is a byproduct of general results about the structure of independent sets in hypergraphs, themselves building on methods of Kleitman and Winston [69] and Sapozhenko [108, 109, 110]. Of course, this summary does a disservice to all three methods, each of which involves the introduction of several new ideas. Surprisingly, all three proofs are substantially different and all three methods have their own particular strengths, some of which we will highlight below.

Rather than focusing on these three methods from the outset, we will further describe the developments leading up to them, explaining how these new results fit into the broader context. This will also allow us to review many of the important subsequent developments. We begin by discussing random analogues of Ramsey-type theorems.

## 2 Ramsey-type theorems in random sets

Ramsey's theorem [93] states that for any graph  $H$  and any natural number  $r$  there exists  $n$  such that any  $r$ -colouring of the edges of the complete graph  $K_n$  on  $n$  vertices contains a monochromatic

copy of  $H$ . The smallest such  $n$  is known as the  $r$ -colour Ramsey number of  $H$  and denoted  $R(H; r)$ . When  $r = 2$ , we simply write this as  $R(H)$  and when  $H = K_t$ , we just write  $R(t)$ . The result of Erdős mentioned in the introduction then says that  $R(t) \geq 2^{t/2}$ , while an upper bound due to Erdős and Szekeres [37] says that  $R(t) \leq 4^t$ . Though there have been lower order improvements to both of these estimates [17, 115], it remains a major open problem to give an exponential improvement to either of them.

Given a graph  $H$  and a natural number  $r$ , we say that a graph  $G$  is  $(H, r)$ -Ramsey if in any  $r$ -colouring of the edges of  $G$  there is guaranteed to be a monochromatic copy of  $H$ . Ramsey's theorem is the statement that  $K_n$  is  $(H, r)$ -Ramsey for  $n$  sufficiently large, while the overall aim of graph Ramsey theory is to decide which graphs are  $(H, r)$ -Ramsey for a given  $H$  and  $r$ . Though coNP-hard in general [15], this problem has borne much fruit and there is now a large theory with many interesting and important results (see, for example, [59]). One of the highlights of this theory is the following random Ramsey theorem of Rödl and Ruciński [96, 97, 98], which determines the threshold for Ramsey's theorem to hold in random graphs. As mentioned in the introduction, this result built on earlier work of Frankl and Rödl [40] and Łuczak, Ruciński and Voigt [86]. Here and throughout the paper, we will write  $v(H)$  and  $e(H)$  for the number of vertices and edges, respectively, of a graph  $H$ .

**Theorem 2.1.** *For any graph  $H$  that is not a forest consisting of stars and paths of length 3 and any positive integer  $r \geq 2$ , there exist positive constants  $c$  and  $C$  such that*

$$\lim_{n \rightarrow \infty} \mathbb{P}[G_{n,p} \text{ is } (H, r)\text{-Ramsey}] = \begin{cases} 0 & \text{if } p < cn^{-1/m_2(H)}, \\ 1 & \text{if } p > Cn^{-1/m_2(H)}, \end{cases}$$

where

$$m_2(H) = \max \left\{ \frac{e(H') - 1}{v(H') - 2} : H' \subseteq H \text{ and } v(H') \geq 3 \right\}.$$

There are two parts to this theorem, one part saying that for  $p < cn^{-1/m_2(H)}$  the random graph  $G_{n,p}$  is highly unlikely to be  $(H, r)$ -Ramsey and the other saying that for  $p > Cn^{-1/m_2(H)}$  it is almost surely  $(H, r)$ -Ramsey. Following standard usage, we will refer to these two parts as the 0-statement and the 1-statement, respectively.

The threshold in Theorem 2.1 occurs at  $p^* = n^{-1/m_2(H)}$ . This is the largest probability for which there is some subgraph  $H'$  of  $H$  such that the number of copies of  $H'$  in  $G_{n,p}$  is approximately the same as the number of edges. For  $p$  significantly smaller than  $p^*$ , the number of copies of  $H'$  will also be significantly smaller than the number of edges. A rather delicate argument [96] then allows one to show that the edges of the graph may be colored so as to avoid any monochromatic copies of  $H'$ . For  $p$  significantly larger than  $p^*$ , almost every edge in the random graph is contained in many copies of every subgraph of  $H$ . The intuition, which takes substantial effort to make rigorous [98], is that these overlaps are enough to force the graph to be Ramsey.

That the proof of the 0-statement is delicate is betrayed by the omitted cases, which have smaller thresholds. For example, if a graph contains the star  $K_{1,r(t-1)+1}$ , then any  $r$ -colouring of the edges of this graph will contain a monochromatic  $K_{1,t}$ . However, the threshold for the appearance of  $K_{1,r(t-1)+1}$  is lower than the threshold suggested by  $m_2(K_{1,t})$ . A more subtle case is when  $H = P_4$ ,

the path with 3 edges (and 4 vertices), and  $r = 2$ . In this case, a cycle of length five with a pendant edge at each vertex is  $(P_4, 2)$ -Ramsey. While the threshold for the appearance of these graphs is at  $n^{-1}$ , which is the same as  $n^{-1/m_2(H)}$ , the threshold is coarse. This means that they start to appear with positive probability already when  $p = c/n$  for any positive  $c$ . This implies that the 0-statement only holds when  $p = o(1/n)$ .

It is worth saying a little about the proof of the 1-statement in Theorem 2.1. We will focus on the case when  $H = K_3$  and  $r = 2$ . The key idea is to write  $G_{n,p}$  as the union of two independent random graphs  $G_{n,p_1}$  and  $G_{n,p_2}$ , chosen so that

$$p = p_1 + p_2 - p_1 p_2 \quad \text{and} \quad p_2 = L p_1$$

for some large constant  $L$ . We first expose the smaller random graph  $G_{n,p_1}$ . With high probability, every colouring of  $G_{n,p_1}$  will contain many monochromatic paths of length 2. If  $p_1$  is a sufficiently large multiple of  $1/\sqrt{n}$ , it is also possible to show that with high probability these monochromatic paths are well distributed. In particular, for any given colouring of  $G_{n,p_1}$ , there are at least  $cn^3$  triangles in the underlying graph  $K_n$  such that there is a path of the same colour, say red, between each pair of vertices in each triangle.

We now expose  $G_{n,p_2}$ . If this graph contains any of the  $cn^3$  triangles described above, we are done, since each edge of this triangle must take the colour blue. Otherwise, together with the red connecting path, we would have a red triangle. By Janson's inequality [66], the probability that  $G_{n,p_2}$  does not contain any of the  $cn^3$  triangles associated to this particular colouring is at most  $2^{-c'p_2n^2}$ , where  $c'$  depends on  $c$ . However, we must remember to account for every possible colouring of  $G_{n,p_1}$ . To do this, we take a union bound. Indeed, since there are at most  $2^{p_1n^2}$  colourings of  $G_{n,p_1}$ , the probability that there exists a colouring such that  $G_{n,p_2}$  does not intersect the associated set of triangles is at most  $2^{p_1n^2} 2^{-c'p_2n^2}$ . If we choose  $L$  sufficiently large, this probability tends to zero, completing the proof.

This method also allowed Rödl and Ruciński to determine the threshold for van der Waerden's theorem to hold in random subsets of the integers. Van der Waerden's theorem [124] states that for any natural numbers  $k$  and  $r$  there exists  $n$  such that any  $r$ -colouring of  $[n] := \{1, 2, \dots, n\}$  contains a monochromatic  $k$ -term arithmetic progression, that is, a monochromatic subset of the form  $\{a, a + d, \dots, a + (k - 1)d\}$ . To state the random version of this theorem, we define  $[n]_p$  to be a random subset of  $[n]$  where each element is chosen independently with probability  $p$ . We also say that a subset  $I$  of the integers is  $(k, r)$ -vdW if in any  $r$ -colouring of the points of  $I$  there is a monochromatic  $k$ -term arithmetic progression. Rödl and Ruciński's random van der Waerden theorem [98, 99] is then as follows.

**Theorem 2.2.** *For any positive integers  $k \geq 3$  and  $r \geq 2$ , there exist positive constants  $c$  and  $C$  such that*

$$\lim_{n \rightarrow \infty} \mathbb{P}[[n]_p \text{ is } (k, r)\text{-vdW}] = \begin{cases} 0 & \text{if } p < cn^{-1/(k-1)}, \\ 1 & \text{if } p > Cn^{-1/(k-1)}. \end{cases}$$

The threshold is again a natural one, since it is the point where we expect that most vertices in  $[n]_p$  will be contained in a constant number of  $k$ -term arithmetic progressions. We will say more about this in the next section when we discuss density theorems.

One question left open by the work of Rödl and Ruciński was whether Theorem 2.1 could be extended to hypergraphs. While some partial progress was made [100, 101], the general problem remained open, not least because of the apparent need to apply a hypergraph analogue of the regularity lemma, something which has only been developed in recent years [55, 88, 104, 120]. Approaches which circumvent hypergraph regularity were developed independently by Friedgut, Rödl and Schacht [45] and by Gowers and the author [24], so that the following generalisation of Theorem 2.1 is now known. We write  $G_{n,p}^{(k)}$  for the random  $k$ -uniform hypergraph on  $n$  vertices, where each edge is chosen independently with probability  $p$ .

**Theorem 2.3.** *For any  $k$ -uniform hypergraph  $H$  and any positive integer  $r \geq 2$ , there exists  $C > 0$  such that*

$$\lim_{n \rightarrow \infty} \mathbb{P}[G_{n,p}^{(k)} \text{ is } (H, r)\text{-Ramsey}] = 1 \text{ if } p > Cn^{-1/m_k(H)},$$

where

$$m_k(H) = \max \left\{ \frac{e(H') - 1}{v(H') - k} : H' \subseteq H \text{ and } v(H') \geq k + 1 \right\}.$$

We note that the approach in [24] applies when  $H$  is strictly  $k$ -balanced, that is, when  $m_k(H) > m_k(H')$  for every subgraph  $H'$  of  $H$ . However, almost all hypergraphs, including the complete hypergraph  $K_t^{(k)}$ , satisfy this requirement. A similar caveat applies to many of the theorems stated in this survey. We will usually make this explicit.

The 0-statement corresponding to Theorem 2.3 was considered by Gugelmann, Person, Steger and Thomas (see [61, 62]). In particular, their results imply the corresponding 0-statement for complete hypergraphs. However, there are again cases where the true threshold is smaller than  $n^{-1/m_k(H)}$ . Indeed, the picture seems to be more complicated than for graphs since there are examples other than the natural generalisations of paths and stars for which the 1-statement may be improved. We refer the reader to [61] for a more complete discussion.

One may also consider the threshold for asymmetric Ramsey properties. We say that a graph  $G$  is  $(H_1, H_2, \dots, H_r)$ -Ramsey if any colouring of the edges of  $G$  with colours  $1, 2, \dots, r$  contains a monochromatic copy of  $H_i$  in colour  $i$  for some  $i \in \{1, 2, \dots, r\}$ . A conjecture of Kohayakawa and Kreuter [71], which generalises Theorem 2.1, says that if  $H_1, H_2, \dots, H_r$  are graphs with  $1 < m_2(H_r) \leq \dots \leq m_2(H_1)$ , then the  $(H_1, H_2, \dots, H_r)$ -Ramsey property has a threshold at  $n^{-1/m_2(H_1, H_2)}$ , where

$$m_2(H_1, H_2) = \max \left\{ \frac{e(H'_1)}{v(H'_1) - 2 + 1/m_2(H_2)} : H'_1 \subseteq H_1 \text{ and } v(H'_1) \geq 3 \right\}.$$

Since the 0-statement fails to hold for certain forests, this statement should be qualified further, but it seems likely to hold for most collections of graphs.

Kohayakawa and Kreuter established the conjecture when  $H_1, H_2, \dots, H_r$  are cycles. As noted in [87], the same method shows that the KLR conjecture (which we discuss in Section 4) would imply the 1-statement of the conjecture when  $H_1$  is strictly 2-balanced, that is, when  $m_2(H_1) > m_2(H'_1)$  for all proper subgraphs  $H'_1$ . Since the KLR conjecture is now an established fact, the following theorem is known to hold (as was noted explicitly by Balogh, Morris and Samotij [6]).

**Theorem 2.4.** *For any graphs  $H_1, H_2, \dots, H_r$  with  $1 < m_2(H_r) \leq \dots \leq m_2(H_1)$  and such that  $H_1$  is strictly 2-balanced, there exists  $C > 0$  such that*

$$\lim_{n \rightarrow \infty} \mathbb{P}[G_{n,p} \text{ is } (H_1, H_2, \dots, H_r)\text{-Ramsey}] = 1 \text{ if } p > Cn^{-1/m_2(H_1, H_2)}.$$

A slightly weaker statement was established by Kohayakawa, Schacht and Spöhel [79] without appealing to the KLR conjecture. Their proof is much closer in spirit to Rödl and Ruciński's proof of Theorem 2.1. A corresponding 0-statement when  $H_1, H_2, \dots, H_r$  are cliques was established by Marciniszyn, Skokan, Spöhel and Steger [87]. However, the 0-statement remains open in general.

The methods developed in [24] and [45] also allow one to extend Rödl and Ruciński's results on random analogues of van der Waerden's theorem to a more general setting. A classical theorem of Rado [92] generalises van der Waerden's theorem by establishing necessary and sufficient conditions for a system of homogeneous linear equations

$$\sum_{j=1}^k a_{ij}x_j = 0 \text{ for } 1 \leq i \leq \ell$$

to be partition regular, that is, to be such that any finite colouring of the natural numbers contains a monochromatic solution  $(x_1, x_2, \dots, x_k)$  to this system of equations. To give an example, the solutions to the system of equations  $x_i + x_{i+2} = 2x_{i+1}$  for  $i = 1, 2, \dots, k-2$  are  $k$ -term arithmetic progressions and so van der Waerden's theorem implies that this system of equations is partition regular. An extension of Theorem 2.2 was proved by Rödl and Ruciński in [99], but their 1-statement only applied to density regular systems of equations (though see also [57]). These are systems of equations, like the system defining  $k$ -term arithmetic progressions, whose solutions sets are closed under translation and dilation.

An extension of this theorem which applies to all partition regular systems of equations was proved by Friedgut, Rödl and Schacht [45]. More precisely, they proved a 1-statement, while the 0-statement had been established earlier by Rödl and Ruciński [99]. Since the details are somewhat technical, we refer the interested reader to [45] for further particulars.

We have already mentioned that the result of Frankl and Rödl [40] may be used to prove that there are  $K_4$ -free graphs which are  $(K_3, 2)$ -Ramsey. This was originally proved by Folkman [38] using a constructive argument. More generally, he proved that for any positive integer  $t$  there is a  $K_{t+1}$ -free graph which is  $(K_t, 2)$ -Ramsey. This beautiful result was subsequently extended to  $r$ -colourings by Nešetřil and Rödl [90, 91].

Once we know that these graphs exist, it is natural to try and estimate their size. We define the Folkman number  $F(t)$  to be the smallest natural number  $n$  such that there exists a  $K_{t+1}$ -free graph  $G$  on  $n$  vertices with the property that every 2-colouring of the edges of  $G$  contains a monochromatic  $K_t$ . The upper bounds on  $F(t)$  which come from the constructive proofs tend to have a dependency on  $t$  which, with a conservative estimate, is at least tower-type, that is, a tower of twos of height at least  $t$ . On the other hand, the lower bound is essentially the same as for Ramsey's theorem, that is,  $F(t) \geq 2^{c^t}$ .

Very recently, it was noted that some of the methods for proving Ramsey-type theorems in random sets yield significantly stronger bounds for Folkman numbers [25, 102]. In particular, the

following result was proved by Rödl, Ruciński and Schacht [102]. Their proof relies heavily on the hypergraph container results developed by Balogh, Morris and Samotij [6] and Saxton and Thomason [111] and an observation of Nenadov and Steger [89] that allows one to apply this machinery in the Ramsey setting.

**Theorem 2.5.** *There exists a constant  $c$  such that*

$$F(t) \leq 2^{ct^4 \log t}.$$

This bound is tantalisingly close to the lower bound and it would be of great interest to improve it further. Since we have now brought our discussions of Ramsey-type theorems in random sets full circle, this provides a convenient departure point to move on to discussing density theorems in random sets, a topic about which much less was known before recent developments.

### 3 Density theorems in random sets

Turán's theorem [123] states that the largest  $K_t$ -free subgraph of  $K_n$  has at most  $\left(1 - \frac{1}{t-1}\right) \frac{n^2}{2}$  edges. Moreover, the unique  $K_t$ -free subgraph achieving this maximum is the  $(t-1)$ -partite graph with vertex sets  $V_1, V_2, \dots, V_{t-1}$ , where each set is of order  $\lfloor \frac{n}{t-1} \rfloor$  or  $\lceil \frac{n}{t-1} \rceil$ . In particular, for  $t = 3$ , the triangle-free subgraph of  $K_n$  with the most edges is a bipartite graph with parts of order  $\lfloor \frac{n}{2} \rfloor$  or  $\lceil \frac{n}{2} \rceil$ . A substantial generalisation of this theorem, known as the Erdős–Stone–Simonovits theorem [34, 36], states that for any graph  $H$  the largest  $H$ -free subgraph of  $K_n$  has at most  $\left(1 - \frac{1}{\chi(H)-1} + o(1)\right) \binom{n}{2}$  edges, where  $\chi(H)$  is the chromatic number of  $H$ .

We say that a graph  $G$  is  $(H, \epsilon)$ -Turán if every subgraph of  $G$  with at least  $\left(1 - \frac{1}{\chi(H)-1} + \epsilon\right) e(G)$  edges contains a copy of  $H$ . The original work of Frankl and Rödl [40] on Ramsey properties in random graphs was actually motivated by a problem of Erdős and Nešetřil concerning an analogue of Folkman's theorem for the  $(H, \epsilon)$ -Turán property. Specifically, they asked whether there exist  $K_4$ -free graphs which are  $(K_3, \epsilon)$ -Turán and Frankl and Rödl showed that there are. Though not stated explicitly in their paper, Frankl and Rödl's method implies that for any  $\epsilon > 0$  there exists  $C > 0$  such that if  $p > C/\sqrt{n}$  then

$$\lim_{n \rightarrow \infty} \mathbb{P}[G_{n,p} \text{ is } (K_3, \epsilon)\text{-Turán}] = 1.$$

Unlike Ramsey properties, the corresponding 0-statement is easy to prove. Indeed, for  $p$  a sufficiently small multiple of  $1/\sqrt{n}$ , the number of triangles in  $G_{n,p}$  will be significantly smaller than the number of edges. We may therefore remove all copies of  $K_3$  by deleting one edge from each copy, leaving a subgraph which is triangle-free but contains at least  $(1 - \delta)e(G_{n,p})$  edges.

A similar argument provides a lower bound for all  $H$ . That is, if the number of copies of  $H$  is significantly smaller than the number of edges, we can remove all copies of  $H$  by deleting one edge from each copy. Therefore, if  $p^{e(H)} n^{v(H)} \ll pn^2$ , that is,  $p \ll n^{-(v(H)-2)/(e(H)-1)}$ , the  $(H, \epsilon)$ -Turán property cannot hold. Since the same argument applies for any subgraph  $H'$  of  $H$ , it is easy to see that for  $p \ll n^{-1/m_2(H)}$  the random graph  $G_{n,p}$  cannot be  $(H, \epsilon)$ -Turán. Here  $m_2(H)$  is defined as

in Theorem 2.1, that is,

$$m_2(H) = \max \left\{ \frac{e(H') - 1}{v(H') - 2} : H' \subseteq H \text{ and } v(H') \geq 3 \right\},$$

The natural conjecture that the  $(H, \epsilon)$ -Turán property holds in random graphs with  $p \gg n^{-1/m_2(H)}$  was first stated by Haxell, Kohayakawa and Łuczak [63, 64] and reiterated by Kohayakawa, Łuczak and Rödl [73].

Until recently, this conjecture was only known to hold for a small collection of graphs, including  $K_3$ ,  $K_4$  and  $K_5$  [40, 73, 53] and all cycles [46, 63, 64] (see also [76, 117]). A verification of the conjecture for all graphs was completed by Schacht [112] and by Gowers and the author [24], although we must qualify this statement by saying that the results of [24] apply when  $H$  is strictly 2-balanced, that is, when  $m_2(H') < m_2(H)$  for all  $H' \subset H$ . However, the class of strictly 2-balanced graphs includes many of the graphs one normally considers, such as cliques and cycles.

**Theorem 3.1.** *For any graph  $H$  and any  $\epsilon > 0$ , there exist positive constants  $c$  and  $C$  such that*

$$\lim_{n \rightarrow \infty} \mathbb{P}[G_{n,p} \text{ is } (H, \epsilon)\text{-Turán}] = \begin{cases} 0 & \text{if } p < cn^{-1/m_2(H)}, \\ 1 & \text{if } p > Cn^{-1/m_2(H)}. \end{cases}$$

As mentioned in the introduction, Schacht's proof of Theorem 3.1 builds on Rödl and Ruciński's proof of Theorem 2.1. In the last section, we gave a brief description of their method, showing how it was best to think of the random graph  $G_{n,p}$  as the union of two independent random graphs  $G_{n,p_1}$  and  $G_{n,p_2}$ . In Schacht's method, this multi-round exposure is taken further, the rough idea being to expose  $G_{n,p}$  over several successive rounds and to apply a density increment argument.

The method employed in [24] relies upon proving a transference principle, an idea which originates in the work of Green and Tao [60] (see also [56, 94]). In the case of triangles, this transference principle says that for  $p \geq C/\sqrt{n}$  any subgraph  $G$  of  $G_{n,p}$  may be modelled by a subgraph  $K$  of the complete graph  $K_n$  in such a way that the proportion of edges and triangles in  $K$  is close to the proportion of edges and triangles in  $G$ . That is, if the sparse graph  $G$  contains  $c_1 pn^2$  edges and  $c_2 p^3 n^3$  triangles, then the dense model  $K$  will contain approximately  $c_1 n^2$  edges and  $c_2 n^3$  triangles.

Suppose now that we wish to prove Turán's theorem for triangles relative to a random graph. Given a subgraph  $G$  of  $G_{n,p}$  with  $(\frac{1}{2} + \epsilon) p \binom{n}{2}$  edges, we know, once our approximation is sufficiently good, that its dense model  $K$  has at least  $(\frac{1}{2} + \frac{\epsilon}{2}) \binom{n}{2}$  edges. A robust version of Turán's theorem [35] then implies that  $K$  contains at least  $cn^3$  triangles for some  $c > 0$  depending on  $\epsilon$ . Provided again that our approximation is sufficiently good, this implies that  $G$  contains at least  $\frac{c}{2} p^3 n^3$  triangles, which is even more than we required.

Though the analogue of Turán's theorem for hypergraphs is rather poorly understood (see, for example, [68]), a similar strategy shows that it is still possible to transfer it to the random setting. To state the result, we need some definitions. Given a  $k$ -uniform hypergraph  $H$ , we let  $\text{ex}(n, H)$  be the largest number of edges in an  $H$ -free subgraph of  $K_n^{(k)}$  and

$$\pi_k(H) = \lim_{n \rightarrow \infty} \frac{\text{ex}(n, H)}{\binom{n}{k}}.$$



We then say that a  $k$ -uniform hypergraph  $G$  is  $(H, \epsilon)$ -Turán if every subgraph of  $G$  with at least  $(\pi_k(H) + \epsilon)e(G)$  edges contains a copy of  $H$ . Let  $m_k(H)$  be defined as in the previous subsection, that is,

$$m_k(H) = \max \left\{ \frac{e(H') - 1}{v(H') - k} : H' \subseteq H \text{ and } v(H') \geq k + 1 \right\}.$$

Then the analogue of Theorem 3.1, proved in [24, 112], states that the property of being  $(H, \epsilon)$ -Turán for a  $k$ -uniform hypergraph  $H$  has a threshold at  $n^{-1/m_k(H)}$ .

**Theorem 3.2.** *For any  $k$ -uniform hypergraph  $H$  and any  $\epsilon > 0$ , there exist positive constants  $c$  and  $C$  such that*

$$\lim_{n \rightarrow \infty} \mathbb{P}[G_{n,p}^{(k)} \text{ is } (H, \epsilon)\text{-Turán}] = \begin{cases} 0 & \text{if } p < cn^{-1/m_k(H)}, \\ 1 & \text{if } p > Cn^{-1/m_k(H)}. \end{cases}$$

One structural counterpart to Turán's theorem is the Erdős-Simonovits stability theorem [114]. This says that for any graph  $H$  with  $\chi(H) \geq 3$  and any  $\epsilon > 0$ , there exists  $\delta > 0$  such that any  $H$ -free subgraph of  $K_n$  with at least  $\left(1 - \frac{1}{\chi(H)-1} - \delta\right) \binom{n}{2}$  edges may be made  $(\chi(H) - 1)$ -partite by removing at most  $\epsilon n^2$  edges. The following sparse analogue of this result was originally proved in [24] for strictly 2-balanced graphs. Later, Samotij [107] found a way to amend Schacht's method so that it applied to stability statements, extending this result to all graphs.

**Theorem 3.3.** *For any graph  $H$  with  $\chi(H) \geq 3$  and any  $\epsilon > 0$ , there exist positive constants  $\delta$  and  $C$  such that if  $p \geq Cn^{-1/m_2(H)}$  the random graph  $G_{n,p}$  a.a.s. has the following property. Every  $H$ -free subgraph of  $G_{n,p}$  with at least  $\left(1 - \frac{1}{\chi(H)-1} - \delta\right) p \binom{n}{2}$  edges can be made  $(\chi(H) - 1)$ -partite by removing at most  $\epsilon p n^2$  edges.*

For cliques, Turán's theorem has a much more precise corresponding structural statement, saying that the largest  $K_t$ -free subgraph is  $(t-1)$ -partite. One may therefore ask when this property holds a.a.s. in the random graph  $G_{n,p}$ . This question was first studied by Babai, Simonovits and Spencer [4] who showed that for  $p > \frac{1}{2}$  the size of the maximum triangle-free subgraph is a.a.s. the same as the size of the largest bipartite subgraph. This result was extended to the range  $p > n^{-c}$  by Brightwell, Panagiotou and Steger [14]. Recently, DeMarco and Kahn [27] proved the following much more precise result.

**Theorem 3.4.** *There is a positive constant  $C$  such that if  $p > C\sqrt{\log n/n}$  then a.a.s. every maximum triangle-free subgraph of  $G_{n,p}$  is bipartite.*

The threshold here is different from the  $1/\sqrt{n}$  we have come to expect. However, the result is sharp up to the constant  $C$ . Indeed, for  $p = 0.1\sqrt{\log n/n}$ , the random graph  $G_{n,p}$  will typically contain a 5-cycle none of whose edges are contained in a triangle. In a forthcoming paper, DeMarco and Kahn [28] prove the following extension of this result to all cliques. Once again, the extra log factors are essential.

**Theorem 3.5.** *For any natural number  $t$ , there exists  $C > 0$  such that if*

$$p > Cn^{-\frac{2}{t+1} \log^{\frac{2}{(t+1)(t-2)}} n}$$

*then a.a.s. every maximum  $K_t$ -free subgraph of  $G_{n,p}$  is  $(t-1)$ -partite.*

We note that a related question, where one wishes to determine the range of  $m$  for which most  $K_t$ -free graphs with  $n$  vertices and  $m$  edges are  $(t-1)$ -partite, was solved recently by Balogh, Morris, Samotij and Warnke [7].

The methods of [24] and [112] also allow one to prove sparse analogues of density statements from other settings. For example, Szemerédi's theorem [118] states that for any natural number  $k$  and any  $\delta > 0$  there exists  $n_0$  such that if  $n \geq n_0$  any subset of  $[n]$  of density at least  $\delta$  contains a  $k$ -term arithmetic progression. This is the density version of van der Waerden's theorem and trivially implies that theorem by taking  $\delta = \frac{1}{r}$  and considering the largest colour class. This theorem and the tools arising in its many proofs [48, 54, 95] have been enormously influential in the development of modern combinatorics.

We say that a subset  $I$  of the integers is  $(k, \delta)$ -Szemerédi if any subset of  $I$  with at least  $\delta|I|$  elements contains an arithmetic progression of length  $k$ . Szemerédi's theorem says that for  $n$  sufficiently large the set  $[n]$  is  $(k, \delta)$ -Szemerédi, while a striking corollary of Green and Tao's work on arithmetic progressions in the primes [60] says that for  $n$  sufficiently large the set of primes up to  $n$  is  $(k, \delta)$ -Szemerédi.

For random subsets of the integers, the  $(k, \delta)$ -Szemerédi property was first studied by Kohayakawa, Łuczak and Rödl [72], who proved that the property of being  $(3, \delta)$ -Szemerédi has a threshold at  $1/\sqrt{n}$ . In general, the natural conjecture is that the  $(k, \delta)$ -Szemerédi property has a threshold at  $n^{-1/(k-1)}$ . The lower bound is again straightforward, since for  $p \ll n^{-1/(k-1)}$  the number of  $k$ -term arithmetic progressions is significantly smaller than the number of elements in the random set  $[n]_p$ , allowing us to remove one element from each arithmetic progression without significantly affecting the density. The corresponding 1-statement was provided in [24] and [112].

**Theorem 3.6.** *For any integer  $k \geq 3$  and  $\delta > 0$ , there exist positive constants  $c$  and  $C$  such that*

$$\lim_{n \rightarrow \infty} \mathbb{P}[[n]_p \text{ is } (k, \delta)\text{-Szemerédi}] = \begin{cases} 0 & \text{if } p < cn^{-1/(k-1)}, \\ 1 & \text{if } p > Cn^{-1/(k-1)}. \end{cases}$$

A particularly satisfying approach to density theorems in random sets is provided by the recent hypergraph containers method of Balogh, Morris and Samotij [6] and Saxton and Thomason [111], the only probabilistic input being Chernoff's inequality and the union bound. In the context of Szemerédi's theorem, one of the main corollaries of this method is the following theorem.

**Theorem 3.7.** *For any integer  $k \geq 3$  and any  $\epsilon > 0$ , there exists  $C > 0$  such that if  $m \geq Cn^{1-1/(k-1)}$ , then there are at most  $\binom{\epsilon n}{m}$  subsets of  $\{1, 2, \dots, n\}$  of order  $m$  which contain no  $k$ -term arithmetic progression.*

Given this statement, which is completely deterministic, it is straightforward to derive the 1-statement in Theorem 3.6, so much so that we may now give the entire calculation. For brevity, we write  $(k, \delta)$ -Sz rather than  $(k, \delta)$ -Szemerédi and  $\mathcal{I}_k(n, \delta pn/2)$  for the collection of subsets of

$\{1, 2, \dots, n\}$  of order  $\delta pn/2$  which contain no  $k$ -term arithmetic progression. We have

$$\begin{aligned}
\mathbb{P}[[n]_p \text{ is not } (k, \delta)\text{-Sz}] &\leq \mathbb{P}[|[n]_p| < pn/2] + \mathbb{P}[|[n]_p| \geq pn/2 \text{ and } [n]_p \text{ is not } (k, \delta)\text{-Sz}] \\
&\leq \exp(-\Omega(pn)) + \mathbb{P}[[n]_p \supseteq I \text{ for some } I \in \mathcal{I}_k(n, \delta pn/2)] \\
&\leq \exp(-\Omega(pn)) + \left(\frac{\epsilon n}{\delta pn/2}\right) p^{\delta pn/2} \\
&\leq \exp(-\Omega(pn)) + \left(\frac{2e\epsilon pn}{\delta pn}\right)^{\delta pn/2} \\
&= \exp(-\Omega(pn)),
\end{aligned}$$

provided  $\epsilon < \delta/2e$ .

Deriving Theorem 3.1 from the results of [6] and [111] involves a little more work. To describe the idea, we focus on the case where  $H = K_3$ . We begin by considering the 3-uniform hypergraph  $\mathcal{G}$  whose vertex set is the collection of edges in  $K_n$  and whose edge set is the collection of triangles in  $K_n$ . Turán's theorem for triangles may then be restated as saying that this 3-uniform hypergraph has no independent set of order greater than  $(\frac{1}{2} + o(1)) |V(\mathcal{G})|$ . We would now like to show that if  $p \geq C/\sqrt{n}$  then the random set  $V(\mathcal{G})_p$  formed by choosing each element of  $V(\mathcal{G})$  independently with probability  $p$  contains no independent set of order greater than  $(\frac{1}{2} + \epsilon) p |V(\mathcal{G})|$ .

One approach would be to use the union bound and Chernoff's inequality to show that with high probability the intersection of the random set with each independent set is as required. An argument of this variety worked in the proof of Theorem 3.6 above, but usually there are far too many independent sets for this approach to be feasible. The main results in both [6] and [111] circumvent this difficulty by showing that there is a substantially smaller collection of almost independent sets which contain all independent sets. Since these sets are almost independent, we know, by the robust version of Turán's theorem, that they must also have size at most  $(\frac{1}{2} + \frac{\epsilon}{2}) |V(\mathcal{G})|$ , say. Applying the union bound over this smaller set then allows us to derive the result.

## 4 Regularity in random graphs

Szemerédi's regularity lemma [119] is one of the cornerstones of modern graph theory (see [81, 103]). Roughly speaking, it says that the vertex set of every graph  $G$  may be divided into a bounded number of parts in such a way that most of the induced bipartite graphs between different parts are pseudorandom. To be more precise, we need some definitions.

We say that a bipartite graph between sets  $U$  and  $V$  is  $\epsilon$ -regular if, for every  $U' \subseteq U$  and  $V' \subseteq V$  with  $|U'| \geq \epsilon|U|$  and  $|V'| \geq \epsilon|V|$ , the density  $d(U', V')$  of edges between  $U'$  and  $V'$  satisfies

$$|d(U', V') - d(U, V)| \leq \epsilon.$$

A partition of the vertex set of a graph into  $t$  pieces  $V_1, \dots, V_t$  is an equipartition if, for every  $1 \leq i, j \leq t$ , we have  $||V_i| - |V_j|| \leq 1$ . Finally, a partition is  $\epsilon$ -regular if it is an equipartition and, for all but at most  $\epsilon t^2$  pairs  $(V_i, V_j)$ , the induced graph between  $V_i$  and  $V_j$  is  $\epsilon$ -regular. Szemerédi's regularity lemma can now be stated as follows.

**Theorem 4.1.** *For any  $\epsilon > 0$ , there exists an integer  $T$  such that every graph  $G$  admits an  $\epsilon$ -regular partition  $V_1, \dots, V_t$  of its vertex set into  $t \leq T$  pieces.*

For sparse graphs – that is, graphs with  $n$  vertices and  $o(n^2)$  edges – the regularity lemma stated above is vacuous, since every equipartition into a bounded number of parts is  $\epsilon$ -regular for  $n$  sufficiently large. However, as observed independently by Kohayakawa [70] and Rödl, there is a meaningful analogue of the regularity lemma for sparse graphs, provided one is willing to restrict consideration to a well-behaved class of graphs.

To make this more precise, we say that a bipartite graph between sets  $U$  and  $V$  is  $(\epsilon, p)$ -regular if, for every  $U' \subseteq U$  and  $V' \subseteq V$  with  $|U'| \geq \epsilon|U|$  and  $|V'| \geq \epsilon|V|$ , the density  $d(U', V')$  of edges between  $U'$  and  $V'$  satisfies

$$|d(U', V') - d(U, V)| \leq \epsilon p.$$

That is, we alter the definition of regularity so that it is relative to a particular density  $p$ , usually chosen to be comparable to the total density between  $U$  and  $V$ . A partition of the vertex set of a graph into  $t$  pieces  $V_1, \dots, V_t$  is then said to be  $(\epsilon, p)$ -regular if it is an equipartition and, for all but at most  $\epsilon t^2$  pairs  $(V_i, V_j)$ , the induced graph between  $V_i$  and  $V_j$  is  $(\epsilon, p)$ -regular.

The class of graphs to which the Kohayakawa-Rödl regularity lemma applies are the so-called upper-uniform graphs [75]. Suppose that  $0 < \eta \leq 1$ ,  $D > 1$  and  $0 < p \leq 1$  are given. We will say that a graph  $G$  is  $(\eta, p, D)$ -upper-uniform if for all disjoint subsets  $U_1$  and  $U_2$  with  $|U_1|, |U_2| \geq \eta|V(G)|$ , the density of edges between  $U_1$  and  $U_2$  satisfies  $d(U_1, U_2) \leq Dp$ . This condition is satisfied for many natural classes of graphs, including all subgraphs of random graphs of density  $p$ . The sparse regularity lemma of Kohayakawa and Rödl is now as follows.

**Theorem 4.2.** *For any  $\epsilon > 0$  and  $D > 1$ , there exists  $\eta > 0$  and an integer  $T$  such that for every  $p \in [0, 1]$ , every graph  $G$  that is  $(\eta, p, D)$ -upper-uniform admits an  $(\epsilon, p)$ -regular partition  $V_1, \dots, V_t$  of its vertex set into  $t \leq T$  pieces.*

A recent variant of this lemma, due to Scott [113], requires no upper-uniformity assumption on  $G$ , although it is often useful to impose such a constraint in practice. Since the two statements are interchangeable when one is dealing with a subgraph of the random graph, we have chosen to describe the original version.

In applications, the regularity method is usually applied in combination with a counting lemma. Roughly speaking, a counting lemma says that if we start with an arbitrary graph  $H$  and replace its vertices by large independent sets and its edges by  $\epsilon$ -regular bipartite graphs with non-negligible density, then this blow-up will contain roughly the expected number of copies of  $H$ . To state this result formally, we again need some definitions.

Given a graph  $H$  with vertex set  $\{1, 2, \dots, k\}$  and a collection of disjoint vertex sets  $V_1, V_2, \dots, V_k$  in a graph  $G$ , we say that a  $k$ -tuple  $(v_1, v_2, \dots, v_k)$  is a canonical copy of  $H$  in  $G$  if  $v_i \in V_i$  for every  $i \in V(H)$  and  $v_i v_j \in E(G)$  for every  $ij \in E(H)$ . We write  $G(H)$  for the number of canonical copies of  $H$  in  $G$ . The counting lemma may now be stated as follows.

**Lemma 4.3.** *For any graph  $H$  with vertex set  $\{1, 2, \dots, k\}$  and any  $\delta > 0$ , there exists a positive constant  $\epsilon$  and an integer  $n_0$  such that the following holds. Let  $n \geq n_0$  and let  $G$  be a graph whose*

vertex set is a disjoint union  $V_1 \cup V_2 \cup \dots \cup V_k$  of sets of size  $n$ . Assume that for each  $ij \in E(H)$ , the bipartite subgraph of  $G$  induced by  $V_i$  and  $V_j$  is  $\epsilon$ -regular and has density  $d_{ij}$ . Then

$$G(H) = \left( \prod_{ij \in E(H)} d_{ij} \pm \delta \right) n^k.$$

When combined with the regularity lemma, this result allows one to prove a number of well-known theorems in extremal graph theory, including the Erdős–Stone–Simonovits theorem, its stability version and the graph removal lemma. In order to extend these results to sparse graphs, one plausible approach, championed by Kohayakawa, Łuczak and Rödl [73], would be to extend Lemma 4.3 to sparse graphs. For example, it would be ideal if we could replace the densities  $d_{ij}$  with  $d_{ij}p$ , the  $\epsilon$ -regularity condition with an  $(\epsilon, p)$ -regularity condition and the conclusion with

$$G(H) = \left( \prod_{ij \in E(H)} d_{ij} \pm \delta \right) p^{e(H)} n^k.$$

We will initially aim for less, only asking to embed a single canonical copy of  $H$ . Unfortunately, for reasons with which we are now familiar, we cannot hope that such a statement holds for small  $p$ . Indeed, if  $p \ll n^{-1/m_2(H)}$ , there is a subgraph  $H'$  of  $H$  for which  $p^{e(H')} n^{v(H')} \ll pn^2$ . We may therefore remove all copies of  $H'$ , and hence  $H$ , from  $G_{n,p}$  while deleting only a small fraction of the edges. The resulting graph is both  $(\epsilon, p)$ -regular, for some small  $\epsilon$ , and  $H$ -free.

Frustratingly, this embedding lemma also fails for larger values of  $p$ . To see this, take a counterexample of the kind just described but with the sets  $V_i$  of order  $r$  for some  $r$  that is much smaller than  $n$ . Now replace each vertex of this small graph by an independent set with  $n/r$  vertices and each edge with a complete bipartite graph. This yields a graph with  $n$  vertices in each  $V_i$ . It is easy to see that the counterexample survives this blowing-up process, implying that the sought-after sparse embedding lemma is false whenever  $p = o(1)$  (see [52, 74]).

However, these counterexamples have a very special structure, an observation that led Kohayakawa, Łuczak and Rödl to conjecture that they might be rare. Roughly speaking, their conjecture, known as the KŁR conjecture, stated that if  $p \gg n^{-1/m_2(H)}$ , then the number of counterexamples to the embedding lemma is so small that  $G_{n,p}$  should not typically contain any such counterexample as a subgraph. Before stating the conjecture (or theorem as it is now), we introduce some notation.

As above, let  $H$  be a graph with vertex set  $\{1, 2, \dots, k\}$ . We denote by  $\mathcal{G}(H, n, m, p, \epsilon)$  the collection of all graphs  $G$  obtained in the following way. The vertex set of  $G$  is a disjoint union  $V_1 \cup V_2 \cup \dots \cup V_k$  of sets of size  $n$ . For each edge  $ij \in E(H)$ , we add an  $(\epsilon, p)$ -regular bipartite graph with  $m$  edges between the pair  $(V_i, V_j)$ . These are the only edges of  $G$ . We also write  $\mathcal{G}^*(H, n, m, p, \epsilon)$  for the set of all  $G \in \mathcal{G}(H, n, m, p, \epsilon)$  that do not contain a canonical copy of  $H$ .

Since the sparse regularity lemma could yield graphs with different densities between the various pairs of vertex sets, it may seem surprising that we are restricting attention to graphs where all the densities are equal. However, it is sufficient to consider just this case. In fact, the KŁR conjecture, which we now state, is more specific still, since it also takes  $p = m/n^2$ . Again, it turns out that from this case one can deduce any other cases that may be needed.

**Theorem 4.4.** *Let  $H$  be a fixed graph and let  $\beta > 0$ . Then there exist positive constants  $C$  and  $\epsilon$  such that*

$$|\mathcal{G}^*(H, n, m, m/n^2, \epsilon)| \leq \beta^m \binom{n^2}{m}^{e(H)}$$

for every  $m \geq Cn^{2-1/m_2(H)}$ .

The KLR conjecture has attracted considerable attention over the past two decades and was resolved for a number of special cases. The cases  $H = K_3$ ,  $K_4$  and  $K_5$  were solved in [72], [51], and [53], respectively. For cycles, the conjecture was proved in [9, 49] (see also [71] for a slightly weaker version). Related results were also given in [50] and [76]. We state it as a theorem because it has now been proved in full generality by Balogh, Morris and Samotij [6] and by Saxton and Thomason [111].

Many of the results discussed in this survey, including Theorems 3.1 and 3.3, follow easily from the KLR conjecture. Indeed, these applications were the original motivation for the conjecture. However, there are situations where an embedding result is not enough: rather than just a single copy of  $H$ , one needs to know that there are many copies. That is, one needs something more like a full counting lemma. Such a counting lemma was provided in a paper of Gowers, Samotij, Schacht and the author [26], the main result of which is the following. We allow for different densities between parts by replacing  $m$  with a vector  $\mathbf{m} = (m_{ij})_{ij \in E(H)}$ .

**Theorem 4.5.** *For any graph  $H$  and any  $\delta, d > 0$ , there exist positive constants  $\epsilon$  and  $\xi$  with the following property. For any  $\eta > 0$ , there is  $C > 0$  such that if  $p \geq Cn^{-1/m_2(H)}$  then a.a.s. the following holds in  $G_{N,p}$ :*

- (i) *For any  $n \geq \eta N$ ,  $\mathbf{m}$  with  $m_{ij} \geq dpn^2$  for all  $ij \in E(H)$  and any subgraph  $G$  of  $G_{N,p}$  in  $\mathcal{G}(H, n, \mathbf{m}, p, \epsilon)$ ,*

$$G(H) \geq \xi \left( \prod_{ij \in E(H)} \frac{m_{ij}}{n^2} \right) n^{v(H)}.$$

- (ii) *Moreover, if  $H$  is strictly 2-balanced, then*

$$G(H) = (1 \pm \delta) \left( \prod_{ij \in E(H)} \frac{m_{ij}}{n^2} \right) n^{v(H)}.$$

We note that Theorem 4.5(i) follows from Samotij's adaptation [107] of Schacht's method [112] (and may also be derived from the work of Saxton and Thomason [111]), while Theorem 4.5(ii) follows from the work of Gowers and the author [24]. Though stronger than Theorem 4.4 in some obvious ways, it is worth noting that Theorem 4.5 does not return the estimate for the number of counterexamples provided by that theorem. This estimate is important for some applications, Theorem 2.4 being a notable example.

One sample application where we need a counting result rather than an embedding result is for proving a random analogue of the graph removal lemma. This theorem, usually attributed to Ruzsa and Szemerédi [106] (though see also [3, 31, 47]), is as follows: for any  $\delta > 0$ , there exists

$\epsilon > 0$  such that if  $G$  is a graph on  $n$  vertices containing at most  $\epsilon n^{v(H)}$  copies of  $H$ , then  $G$  may be made  $H$ -free by deleting at most  $\delta n^2$  edges. Though simple in appearance, this result is surprisingly difficult to prove (see, for example, [18, 39]). It also has some striking consequences, including the  $k = 3$  case of Szemerédi's theorem, originally due to Roth [105]. A sparse random version of the graph removal lemma was conjectured by Łuczak in [85] and proved, for strictly 2-balanced  $H$ , in [24]. The following statement, which applies for all  $H$ , may be found in [26].

**Theorem 4.6.** *For any graph  $H$  and any  $\delta > 0$ , there exist positive constants  $\epsilon$  and  $C$  such that if  $p \geq Cn^{-1/m_2(H)}$  then the following holds a.a.s. in  $G_{n,p}$ . Every subgraph of  $G_{n,p}$  which contains at most  $\epsilon p^{e(H)} n^{v(H)}$  copies of  $H$  may be made  $H$ -free by removing at most  $\delta pn^2$  edges.*

Note that if  $p \leq cn^{-1/m_2(H)}$ , for  $c$  sufficiently small, this statement is trivially true. Indeed, in this range, there exists a subgraph  $H'$  of  $H$  such that the number of copies of  $H'$  in  $G_{n,p}$  is smaller than  $\delta pn^2$ , so we can simply remove one edge from each copy of  $H'$ . One might then conjecture, as Łuczak did, that Theorem 4.6 holds for all values of  $p$ . For 2-balanced graphs, those with  $m_2(H') \leq m_2(H)$  for all  $H' \subset H$ , we may verify this conjecture by taking  $\epsilon$  to be sufficiently small in terms of  $C$ ,  $\delta$ , and  $H$ . For  $p \leq Cn^{-1/m_2(H)}$  and  $\epsilon < \delta C^{-e(H)}$ , the number of copies of  $H$  is at most  $\epsilon p^{e(H)} n^{v(H)} \leq \epsilon C^{e(H)} pn^2 < \delta pn^2$ . Deleting an edge from each copy yields the result.

## 5 Further directions

### 5.1 Sharp thresholds for Ramsey properties

A graph property  $\mathcal{P}$  is said to be monotone if it is closed under the addition of edges, that is,  $G \in \mathcal{P}$  and  $G \subset G'$  implies that  $G' \in \mathcal{P}$ . A result of Bollobás and Thomason [11] shows that any monotone property has a threshold. For example, since Ramsey properties are clearly monotone, this immediately implies that the  $(H, r)$ -Ramsey property and the  $(k, r)$ -vdW property, both defined in Section 2, have thresholds.

Once we have proved that a given property has a threshold, it is often interesting to study this threshold more closely. We say that  $\mathcal{P}$  has a sharp threshold at  $p^* := p^*(n)$  if, for any  $\epsilon > 0$ ,

$$\lim_{n \rightarrow \infty} \mathbb{P}[G_{n,p} \text{ is in } \mathcal{P}] = \begin{cases} 0 & \text{if } p < (1 - \epsilon)p^*, \\ 1 & \text{if } p > (1 + \epsilon)p^*. \end{cases}$$

For example, the properties of being connected and having a Hamiltonian cycle have sharp thresholds, while the property of containing a particular graph  $H$  has a non-sharp or coarse threshold.

A seminal result of Friedgut [41] gives a criterion for assessing whether a monotone property has a sharp threshold or not. Roughly speaking, this criterion says that if the property is globally determined the threshold is sharp, while if it is locally determined it is not. This fits in well with the examples given above, since connectedness and Hamiltonicity are clearly global properties, while having a single copy of a particular  $H$  is decidedly local.

The question of whether Ramsey properties have sharp thresholds was first studied by Friedgut and Krivelevich [43]. They proved, amongst other things, that the  $(H, r)$ -Ramsey property is sharp when  $H$  is any tree other than a star or a path of length three. However, the first substantial

breakthrough was made by Friedgut, Rödl, Ruciński and Tetali [44], who proved that the  $(K_3, 2)$ -Ramsey property has a sharp threshold. Their result may be stated as follows.

**Theorem 5.1.** *There exists a bounded function  $\hat{c} := \hat{c}(n)$  such that for any  $\epsilon > 0$ ,*

$$\lim_{n \rightarrow \infty} \mathbb{P}[G_{n,p} \text{ is } (K_3, 2)\text{-Ramsey}] = \begin{cases} 0 & \text{if } p < (1 - \epsilon)\hat{c}/\sqrt{n}, \\ 1 & \text{if } p > (1 + \epsilon)\hat{c}/\sqrt{n}. \end{cases}$$

A close look at this result reveals an unusual feature: though we know that the threshold is sharp, we do not know exactly where it lies. In principle, the function  $\hat{c}(n)$  could depend on  $n$  and wander up and down between constants  $c$  and  $C$ . However, we expect that the true behaviour should be that it tends towards a constant. It would be very interesting to prove that this is the case. It would also be of great interest to extend Theorem 5.1 to other graphs and a higher number of colours.

More recently, Friedgut, Hàn, Person and Schacht [42] proved that there is a sharp threshold for the appearance of  $k$ -term arithmetic progressions in every 2-colouring of  $[n]_p$ . That is, they showed that the  $(k, 2)$ -vdW property has a sharp threshold. Their proof relies in a fundamental way on the hypergraph containers results discussed throughout this survey.

**Theorem 5.2.** *For every integer  $k \geq 3$ , there exists a bounded function  $\hat{c}_k := \hat{c}_k(n)$  such that for any  $\epsilon > 0$ ,*

$$\lim_{n \rightarrow \infty} \mathbb{P}[G_{n,p} \text{ is } (k, 2)\text{-vdW}] = \begin{cases} 0 & \text{if } p < (1 - \epsilon)\hat{c}_k n^{-1/(k-1)}, \\ 1 & \text{if } p > (1 + \epsilon)\hat{c}_k n^{-1/(k-1)}. \end{cases}$$

It would again be interesting to determine the asymptotic behaviour of  $\hat{c}_k(n)$  or to extend this result to a higher number of colours.

## 5.2 Large subgraph theorems in random graphs

One of the most active areas of research in extremal combinatorics is in finding conditions under which a graph contains certain large or even spanning sparse subgraphs (see, for example, [83]). It is therefore natural to ask whether these results also have random analogues.

One of the standard examples in this area is Dirac's theorem [29], which says that if a graph on  $n$  vertices has minimum degree at least  $n/2$  then it contains a Hamiltonian cycle, that is, a cycle which meets every vertex. The study of random analogues of Dirac's theorem was initiated by Sudakov and Vu [116] and the state of the art is now the following result of Lee and Sudakov [84].

**Theorem 5.3.** *For any  $\epsilon > 0$ , there exists  $C > 0$  such that if  $p \geq C \frac{\log n}{n}$  then a.a.s. every subgraph of  $G_{n,p}$  with minimum degree at least  $(\frac{1}{2} + \epsilon)pn$  contains a Hamiltonian cycle.*

There has also been considerable work on studying random analogues of the bandwidth theorem of Böttcher, Schacht and Taraz [13]. The bandwidth of a graph  $G$  is the smallest  $b$  for which there is an ordering  $v_1, v_2, \dots, v_n$  of the vertices of  $G$  such that  $|i - j| \leq b$  for all edges  $v_i v_j$ . The theorem then states that for any positive integers  $r$  and  $\Delta$  and any  $\gamma > 0$ , there exists an integer  $n_0$  and  $\beta > 0$  such that if  $n \geq n_0$  and  $H$  is an  $n$ -vertex graph with chromatic number  $r$ , maximum



degree  $\Delta$  and bandwidth at most  $\beta n$ , then any graph on  $n$  vertices with minimum degree at least  $(1 - \frac{1}{r} + \gamma)n$  contains a copy of  $H$ .

For the  $r = 2$  case, that is, for bipartite  $H$ , the following random analogue of this theorem was proved by Böttcher, Kohayakawa and Taraz [12].

**Theorem 5.4.** *For any integer  $\Delta \geq 2$  and any  $\eta, \gamma > 0$ , there exist positive constants  $\beta$  and  $C$  such that if  $p \geq C(\log n/n)^{1/\Delta}$  the random graph  $G_{n,p}$  a.a.s. has the following property. Any subgraph of  $G_{n,p}$  with minimum degree at least  $(\frac{1}{2} + \gamma)pn$  contains any bipartite graph on at most  $(1 - \eta)n$  vertices with maximum degree  $\Delta$  and bandwidth at most  $\beta n$ .*

Related results were also proved by Huang, Lee and Sudakov [65]. In particular, they showed that if  $H$  is an  $r$ -partite graph on  $n$  vertices such that every vertex is contained in a triangle, then there exist subgraphs of the random graph  $G_{n,p}$  with minimum degree at least  $(1 - \frac{1}{r} + \gamma)pn$  such that at least  $cp^{-2}$  vertices are not contained in a copy of  $H$ . That is, we cannot hope to cover all vertices when considering random analogues of the bandwidth theorem. However, as suggested by results in [5] and [65], it may still be possible to embed graphs with as many as  $n - Cp^{-2}$  vertices.

A celebrated result of Chvátal, Rödl, Szemerédi and Trotter [16] (see also [19, 58]) states that for any positive integers  $\Delta$  and  $r$ , there exists  $C > 0$  such that if  $H$  is any graph with  $n$  vertices and maximum degree  $\Delta$ , then  $R(H; r) \leq Cn$ . That is, the Ramsey number of bounded degree graphs grows linearly in the number of vertices. However, one can do even better.

Given a graph  $H$  and a natural number  $r$ , we define the size-Ramsey number  $\hat{R}(H; r)$  to be the smallest number of edges in an  $(H, r)$ -Ramsey graph. So we are now interested in minimising the number of edges rather than the number of vertices. A striking result of Beck [8] says that  $\hat{R}(P_n; r) \leq Cn$  for some  $C$  depending only on  $r$ . Using random graphs, Kohayakawa, Rödl, Schacht and Szemerédi [78] recently proved that if  $H$  is any graph with  $n$  vertices and maximum degree  $\Delta$ , then  $\hat{R}(H; r) \leq n^{2 - \frac{1}{\Delta} + o(1)}$ . That is, the size-Ramsey number of bounded degree graphs is subquadratic in the number of vertices. Precisely stated, their main result is the following.

**Theorem 5.5.** *For any integers  $\Delta \geq 2$  and  $r \geq 2$ , there exists  $C > 0$  such that if  $p \geq C(\log N/N)^{1/\Delta}$  the random graph  $G_{N,p}$  with  $N = Cn$  a.a.s. has the following property. Any  $r$ -colouring of the edges of  $G_{N,p}$  contains a colour class which contains every graph on  $n$  vertices with maximum degree  $\Delta$ .*

In a forthcoming paper, Allen, Böttcher, Hàn, Kohayakawa and Person [1] prove a sparse random version of the blow-up lemma. For dense graphs, this result, proved by Komlós, Sárközy and Szemerédi [80], is a standard tool for embedding spanning subgraphs. Its sparse counterpart should allow one to reprove many of the results mentioned in this section in a unified way.

### 5.3 Combinatorial theorems relative to a pseudorandom set

While this survey has focused on combinatorial theorems relative to random sets, analogous questions may also be asked for pseudorandom sets. Much of the work in this direction has focused on the combinatorial properties of the class of  $(p, \beta)$ -jumbled graphs. These graphs, introduced by Thomason [121, 122], have the property that if  $X$  and  $Y$  are vertex subsets, then

$$|e(X, Y) - p|X||Y|| \leq \beta \sqrt{|X||Y|}.$$

As one would expect of a pseudorandom property, the random graph  $G_{n,p}$  is itself  $(p, \beta)$ -jumbled. In this case, with high probability, we may take  $\beta$  to be  $O(\sqrt{pn})$ . This is essentially optimal, that is, there are no  $(p, \beta)$ -jumbled graphs with  $\beta = o(\sqrt{pn})$ . An explicit example of a jumbled graph is the Paley graph. This is the graph with vertex set  $\mathbb{Z}_p$ , where  $p$  is a prime of the form  $4k + 1$ , and edge set given by joining  $x$  and  $y$  if and only if their difference is a quadratic residue. This graph is again optimally jumbled with  $p = \frac{1}{2}$  and  $\beta = O(\sqrt{n})$ . For many more examples, we refer the reader to the survey [82].

For  $(p, \beta)$ -jumbled graphs, one is usually interested in questions of the following form: given a graph property  $\mathcal{P}$ , an integer  $n$  and a density  $p$ , for what values of  $\beta$  is it the case that a  $(p, \beta)$ -jumbled graph on  $n$  vertices satisfies  $\mathcal{P}$ ? To give an example, for any integer  $t \geq 3$ , there exists  $c > 0$  such that if  $\beta \leq cp^{t-1}n$  then any  $(p, \beta)$ -jumbled graph on  $n$  vertices contains a copy of  $K_t$ . For  $t = 3$ , this condition is known to be tight, as shown by an example of Alon [2].

Very recently, a general method for transferring combinatorial theorems to pseudorandom graphs was found by Fox, Zhao and the author [20]. Though we will not attempt an exhaustive survey, the following sample result is representative.

**Theorem 5.6.** *For any integer  $t$  and any  $\epsilon > 0$ , there exist positive constants  $\delta$  and  $c$  such that if  $\beta \leq cp^t n$  then any  $(p, \beta)$ -jumbled graph  $G$  on  $n$  vertices has the following property. Any subgraph of  $G$  containing at most  $\delta p^{\binom{t}{2}} n^t$  copies of  $K_t$  may be made  $K_t$ -free by deleting at most  $\epsilon p n^2$  edges.*

That is, we have an extension of the removal lemma to subgraphs of pseudorandom graphs. Although we have only stated this result for cliques, there is also a more general statement that applies to all graphs. Moreover, with similar conditions on  $\beta$ , it is possible to prove analogues of many different combinatorial statements. For example, the  $(K_t, r)$ -Ramsey property and  $(K_t, \epsilon)$ -Turán property both hold in pseudorandom graphs with  $\beta \leq cp^t n$ .

Unfortunately, there is still a gap in these results, even for triangles. For  $t = 3$ , Theorem 5.6 (which in this case was first proved by Kohayakawa, Rödl, Schacht and Skokan [77]) says that if  $\beta \leq cp^3 n$  then the triangle removal lemma holds for subgraphs of a  $(p, \beta)$ -jumbled graph on  $n$  vertices. However, it may well be the case that  $\beta \leq cp^2 n$  is sufficient. If true, Alon's example would imply that such a result was optimal.

The method of [20] was extended to hypergraphs in [21], under a different type of pseudorandomness hypothesis (though see also [22]). This result was then used to prove a pseudorandom analogue of Szemerédi's theorem. Such a result was a key ingredient in Green and Tao's proof that the primes contain arbitrarily long arithmetic progressions. Their original result states that if a subset of the integers satisfies two pseudorandomness conditions, the linear forms condition and the correlation condition, then it is  $(k, \delta)$ -Szemerédi. Our results allow one to remove the correlation condition from this statement. Due to space constraints, we are unable to say more here. However, we refer the reader to [23] for further details.

**Acknowledgements.** I would like to thank Jacob Fox, Tim Gowers, Rob Morris, Wojtek Samotij, Mathias Schacht, Benny Sudakov and Yufei Zhao for helpful discussions on the topics in this paper. Particular thanks are due to Wojtek Samotij for a number of detailed comments on the manuscript.

## References

- [1] P. Allen, J. Böttcher, H. Hàn, Y. Kohayakawa and Y. Person, Blow-up lemmas for sparse graphs, *preprint*.
- [2] N. Alon, Explicit Ramsey graphs and orthonormal labellings, *Electron. J. Combin.* **1** (1994), Research paper 12, 8pp.
- [3] N. Alon, R. A. Duke, H. Lefmann, V. Rödl and R. Yuster, The algorithmic aspects of the regularity lemma, *J. Algorithms* **16** (1994), 80–109.
- [4] L. Babai, M. Simonovits and J. Spencer, Extremal subgraphs of random graphs, *J. Graph Theory* **14** (1990), 599–622.
- [5] J. Balogh, C. Lee and W. Samotij, Corrádi and Hajnal’s theorem for sparse random graphs, *Combin. Probab. Comput.* **21** (2012), 23–55.
- [6] J. Balogh, R. Morris and W. Samotij, Independent sets in hypergraphs, *preprint*.
- [7] J. Balogh, R. Morris, W. Samotij and L. Warnke, The typical structure of sparse  $K_{r+1}$ -free graphs, *preprint*.
- [8] J. Beck, On size Ramsey numbers of paths, trees and circuits I, *J. Graph Theory* **7** (1983), 115–129.
- [9] M. Behrisch, Random graphs without a short cycle, Master’s thesis, Humboldt Universität zu Berlin, 2002.
- [10] B. Bollobás, **Random graphs**, second edition, Cambridge Studies in Advanced Mathematics 73, Cambridge University Press, Cambridge, 2001.
- [11] B. Bollobás and A. G. Thomason, Threshold functions, *Combinatorica* **7** (1987), 35–38.
- [12] J. Böttcher, Y. Kohayakawa and A. Taraz, Almost spanning subgraphs of random graphs after adversarial edge removal, *Combin. Probab. Comput.* **22** (2013), 639–683.
- [13] J. Böttcher, M. Schacht and A. Taraz, Proof of the bandwidth conjecture of Bollobás and Komlós, *Math. Ann.* **343** (2009), 175–205.
- [14] G. Brightwell, K. Panagiotou and A. Steger, Extremal subgraphs of random graphs, *Random Structures Algorithms* **41** (2012), 147–178.
- [15] S. Burr, On the computational complexity of Ramsey-type problems, in Mathematics of Ramsey theory, 46–52, Algorithms Combin., 5, Springer, Berlin, 1990.
- [16] V. Chvátal, V. Rödl, E. Szemerédi and W. T. Trotter Jr, The Ramsey number of a graph with bounded maximum degree, *J. Combin. Theory Ser. B* **34** (1983), 239–243.

- [17] D. Conlon, A new upper bound for diagonal Ramsey numbers, *Ann. of Math.* **170** (2009), 941–960.
- [18] D. Conlon and J. Fox, Graph removal lemmas, in *Surveys in Combinatorics 2013*, London Math. Soc. Lecture Note Ser., Vol. 409, 1–50, Cambridge University Press, Cambridge, 2013.
- [19] D. Conlon, J. Fox and B. Sudakov, On two problems in graph Ramsey theory, *Combinatorica* **32** (2012), 513–535.
- [20] D. Conlon, J. Fox and Y. Zhao, Extremal results in sparse pseudorandom graphs, *Adv. Math.* **256** (2014), 206–290.
- [21] D. Conlon, J. Fox and Y. Zhao, A relative Szemerédi theorem, *submitted*.
- [22] D. Conlon, J. Fox and Y. Zhao, Linear forms from the Gowers uniformity norm, *unpublished note*.
- [23] D. Conlon, J. Fox and Y. Zhao, The Green-Tao theorem: an exposition, *submitted*.
- [24] D. Conlon and W. T. Gowers, Combinatorial theorems in sparse random sets, *submitted*.
- [25] D. Conlon and W. T. Gowers, An upper bound for Folkman numbers, *submitted*.
- [26] D. Conlon, W. T. Gowers, W. Samotij and M. Schacht, On the KLR conjecture in random graphs, to appear in *Israel J. Math.*
- [27] B. DeMarco and J. Kahn, Mantel’s theorem in random graphs, *preprint*.
- [28] B. DeMarco and J. Kahn, Turán’s theorem in random graphs, *preprint*.
- [29] G. A. Dirac, Some theorems on abstract graphs, *Proc. London Math. Soc.* **2** (1952), 69–81.
- [30] P. Erdős, Some remarks on the theory of graphs, *Bull. Amer. Math. Soc.* **53** (1947), 292–294.
- [31] P. Erdős, P. Frankl and V. Rödl, The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent, *Graphs Combin.* **2** (1986), 113–121.
- [32] P. Erdős and A. Rényi, On random graphs I, *Publ. Math. Debrecen* **6** (1959), 290–297.
- [33] P. Erdős and A. Rényi, On the evolution of random graphs, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **5** (1960), 17–61.
- [34] P. Erdős and M. Simonovits, A limit theorem in graph theory, *Studia Sci. Math. Hungar.* **1** (1966), 51–57.
- [35] P. Erdős and M. Simonovits, Supersaturated graphs and hypergraphs, *Combinatorica* **3** (1983), 181–192.

- [36] P. Erdős and A. H. Stone, On the structure of linear graphs, *Bull. Amer. Math. Soc.* **52** (1946), 1087–1091.
- [37] P. Erdős and G. Szekeres, A combinatorial problem in geometry, *Compos. Math.* **2** (1935), 463–470.
- [38] J. Folkman, Graphs with monochromatic complete subgraphs in every edge coloring, *SIAM J. Appl. Math.* **18** (1970), 19–24.
- [39] J. Fox, A new proof of the graph removal lemma, *Ann. of Math.* **174** (2011), 561–579.
- [40] P. Frankl and V. Rödl, Large triangle-free subgraphs in graphs without  $K_4$ , *Graphs Combin.* **2** (1986), 135–144.
- [41] E. Friedgut, Sharp thresholds of graph properties, and the k-sat problem, *J. Amer. Math. Soc.* **12** (1999), 1017–1054.
- [42] E. Friedgut, H. Hàn, Y. Person and M. Schacht, A sharp threshold for van der Waerden’s theorem in random subsets of  $\mathbb{Z}/n\mathbb{Z}$ , *preprint*.
- [43] E. Friedgut and M. Krivelevich, Sharp thresholds for Ramsey properties of random graphs, *Random Structures Algorithms* **17** (2000), 1–19.
- [44] E. Friedgut, V. Rödl, A. Ruciński and P. Tetali, A sharp threshold for random graphs with a monochromatic triangle in every edge coloring, *Mem. Amer. Math. Soc.* **179** (2006), no. 845, vi + 66 pp.
- [45] E. Friedgut, V. Rödl and M. Schacht, Ramsey properties of discrete random structures, *Random Structures Algorithms* **37** (2010), 407–436.
- [46] Z. Füredi, Random Ramsey graphs for the four-cycle, *Discrete Math.* **126** (1994), 407–410.
- [47] Z. Füredi, Extremal hypergraphs and combinatorial geometry, in Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Zürich, 1994), 1343–1352, Birkhäuser, Basel, 1995.
- [48] H. Furstenberg, Ergodic behaviour of diagonal measures and a theorem of Szemerédi on arithmetic progressions, *J. Analyse Math.* **31** (1977), 204–256.
- [49] S. Gerke, Y. Kohayakawa, V. Rödl and A. Steger, Small subsets inherit sparse  $\epsilon$ -regularity, *J. Combin. Theory Ser. B* **97** (2007), 34–56.
- [50] S. Gerke, M. Marciniszyn and A. Steger, A probabilistic counting lemma for complete graphs, *Random Structures Algorithms*, **31** (2007), 517–534.
- [51] S. Gerke, H. J. Prömel, T. Schickinger, A. Steger and A. Taraz,  $K_4$ -free subgraphs of random graphs revisited, *Combinatorica* **27** (2007), 329–365.

- [52] S. Gerke and A. Steger, The sparse regularity lemma and its applications, in *Surveys in Combinatorics 2005*, London Math. Soc. Lecture Note Ser., Vol. 327, 227–258, Cambridge University Press, Cambridge, 2005.
- [53] S. Gerke, A. Steger and T. Schickinger,  $K_5$ -free subgraphs of random graphs, *Random Structures Algorithms* **24** (2004), 194–232.
- [54] W. T. Gowers, A new proof of Szemerédi’s theorem, *Geom. Funct. Anal.* **11** (2001), 465–588.
- [55] W. T. Gowers, Hypergraph regularity and the multidimensional Szemerédi theorem, *Ann. of Math.* **166** (2007), 897–946.
- [56] W. T. Gowers, Decompositions, approximate structure, transference, and the Hahn-Banach theorem, *Bull. London Math. Soc.* **42** (2010), 573–606.
- [57] R. L. Graham, V. Rödl and A. Ruciński, On Schur properties of random subsets of integers, *J. Number Theory* **61** (1996), 388–408.
- [58] R. L. Graham, V. Rödl and A. Ruciński, On graphs with linear Ramsey numbers, *J. Graph Theory* **35** (2000), 176–192.
- [59] R. L. Graham, B. L. Rothschild and J. H. Spencer, **Ramsey theory**, second edition, John Wiley & Sons, 1990.
- [60] B. Green and T. Tao, The primes contain arbitrarily long arithmetic progressions, *Ann. of Math.* **167** (2008), 481–547.
- [61] L. Gugelmann, Ramsey properties of random graphs and hypergraphs, PhD thesis, ETH Zürich, 2013.
- [62] L. Gugelmann, Y. Person, A. Steger and H. Thomas, A randomized version of Ramsey’s theorem, *Random Structures Algorithms* **41** (2012), 488–505.
- [63] P. E. Haxell, Y. Kohayakawa and T. Łuczak, Turán’s extremal problem in random graphs: forbidding even cycles, *J. Combin. Theory Ser. B* **64** (1995), 273–287.
- [64] P. E. Haxell, Y. Kohayakawa and T. Łuczak, Turán’s extremal problem in random graphs: forbidding odd cycles, *Combinatorica* **16** (1996), 107–122.
- [65] H. Huang, C. Lee and B. Sudakov, Bandwidth theorem for random graphs, *J. Combin. Theory Ser. B* **102** (2012), 14–37.
- [66] S. Janson, Poisson approximation for large deviations, *Random Structures Algorithms* **1** (1990), 221–229.
- [67] S. Janson, T. Łuczak and A. Ruciński, **Random graphs**, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2000.

- [68] P. Keevash, Hypergraph Turán Problems, in Surveys in Combinatorics 2011, London Math. Soc. Lecture Note Ser., Vol. 392, 83–140, Cambridge University Press, Cambridge, 2011.
- [69] D. J. Kleitman and K. J. Winston, On the number of graphs without 4-cycles, *Discrete Math.* **41** (1982), 167–172.
- [70] Y. Kohayakawa, Szemerédi’s regularity lemma for sparse graphs, in Foundations of computational mathematics (Rio de Janeiro, 1997), 216–230, Springer, Berlin, 1997.
- [71] Y. Kohayakawa and B. Kreuter, Threshold functions for asymmetric Ramsey properties involving cycles, *Random Structures Algorithms* **11** (1997), 245–276.
- [72] Y. Kohayakawa, T. Łuczak and V. Rödl, Arithmetic progressions of length three in subsets of a random set, *Acta Arith.* **75** (1996), 133–163.
- [73] Y. Kohayakawa, T. Łuczak and V. Rödl, On  $K_4$ -free subgraphs of random graphs, *Combinatorica* **17** (1997), 173–213.
- [74] Y. Kohayakawa and V. Rödl, Regular pairs in sparse random graphs I, *Random Structures Algorithms* **22** (2003), 359–434.
- [75] Y. Kohayakawa and V. Rödl, Szemerédi’s regularity lemma and quasi-randomness, in Recent advances in algorithms and combinatorics, CMS Books Math./Ouvrages Math. SMC, 11, 289–351, Springer, New York, 2003.
- [76] Y. Kohayakawa, V. Rödl and M. Schacht, The Turán theorem for random graphs, *Combin. Probab. Comput.* **13** (2004), 61–91.
- [77] Y. Kohayakawa, V. Rödl, M. Schacht and J. Skokan, On the triangle removal lemma for subgraphs of sparse pseudorandom graphs, in An irregular mind (Szemerédi is 70), Bolyai Soc. Math. Stud., Vol. 21, 359–404, Springer, Berlin, 2010.
- [78] Y. Kohayakawa, V. Rödl, M. Schacht and E. Szemerédi, Sparse partition universal graphs for graphs of bounded degree, *Adv. Math.* **226** (2011), 5041–5065.
- [79] Y. Kohayakawa, M. Schacht and R. Spöhel, Upper bounds on probability thresholds for asymmetric Ramsey properties, *Random Structures Algorithms* **44** (2014), 1–28.
- [80] J. Komlós, G. Sárközy and E. Szemerédi, Blow-up lemma, *Combinatorica* **17** (1997), 109–123.
- [81] J. Komlós, A. Shokoufandeh, M. Simonovits and E. Szemerédi, The regularity lemma and its applications in graph theory, in Theoretical aspects of computer science (Tehran, 2000), Lecture Notes in Comput. Sci., Vol. 2292, 84–112, Springer, Berlin, 2002.
- [82] M. Krivelevich and B. Sudakov, Pseudo-random graphs, in More sets, graphs and numbers, Bolyai Soc. Math. Stud., Vol. 15, 199–262, Springer, Berlin, 2006.

- [83] D. Kühn and D. Osthus, Embedding large subgraphs into dense graphs, in *Surveys in Combinatorics 2009*, London Math. Soc. Lecture Note Ser., Vol. 365, 137–167, Cambridge University Press, Cambridge, 2009.
- [84] C. Lee and B. Sudakov, Dirac’s theorem for random graphs, *Random Structures Algorithms* **41** (2012), 293–305.
- [85] T. Łuczak, Randomness and regularity, in *International Congress of Mathematicians*, Vol. III, 899–909, Eur. Math. Soc., Zürich, 2006.
- [86] T. Łuczak, A. Ruciński and B. Voigt, Ramsey properties of random graphs, *J. Combin. Theory Ser. B* **56** (1992), 55–68.
- [87] M. Marciniszyn, J. Skokan, R. Spöhel and A. Steger, Asymmetric Ramsey properties of random graphs involving cliques, *Random Structures Algorithms* **34** (2009), 419–453.
- [88] B. Nagle, V. Rödl and M. Schacht, The counting lemma for regular  $k$ -uniform hypergraphs, *Random Structures Algorithms* **28** (2006), 113–179.
- [89] R. Nenadov and A. Steger, A short proof of the random Ramsey theorem, to appear in *Combin. Probab. Comput.*
- [90] J. Nešetřil and V. Rödl, The Ramsey property for graphs with forbidden complete subgraphs, *J. Combin. Theory Ser. B* **20** (1976), 243–249.
- [91] J. Nešetřil and V. Rödl, Simple proof of the existence of restricted Ramsey graphs by means of a partite construction, *Combinatorica* **1** (1981), 199–202.
- [92] R. Rado, Note on combinatorial analysis, *Proc. London Math. Soc.* **48** (1941), 122–160.
- [93] F. P. Ramsey, On a problem of formal logic, *Proc. London Math. Soc.* **30** (1930), 264–286.
- [94] Omer Reingold, Luca Trevisan, Madhur Tulsiani and Salil Vadhan, Dense Subsets of Pseudorandom Sets, in *Proceedings of the 2008 49th Annual IEEE Symposium on Foundations of Computer Science*, 76–85, IEEE Computer Society, Washington DC, 2008.
- [95] V. Rödl, Quasi-randomness and the regularity method in hypergraphs, *preprint*.
- [96] V. Rödl and A. Ruciński, Lower bounds on probability thresholds for Ramsey properties, in *Combinatorics, Paul Erdős is Eighty*, Vol. 1, 317–346, Bolyai Soc. Math. Studies, János Bolyai Math. Soc., Budapest, 1993.
- [97] V. Rödl and A. Ruciński, Random graphs with monochromatic triangles in every edge coloring, *Random Structures Algorithms* **5** (1994), 253–270.
- [98] V. Rödl and A. Ruciński, Threshold functions for Ramsey properties, *J. Amer. Math. Soc.* **8** (1995), 917–942.



- [99] V. Rödl and A. Ruciński, Rado partition theorem for random subsets of integers, *Proc. London Math. Soc.* **74** (1997), 481–502.
- [100] V. Rödl and A. Ruciński, Ramsey properties of random hypergraphs, *J. Combin. Theory Ser. A* **81** (1998), 1–33.
- [101] V. Rödl, A. Ruciński and M. Schacht, Ramsey properties of random  $k$ -partite,  $k$ -uniform hypergraphs, *SIAM J. Discrete Math.* **21** (2007), 442–460.
- [102] V. Rödl, A. Ruciński and M. Schacht, Ramsey properties of random graphs and Folkman numbers, *preprint*.
- [103] V. Rödl and M. Schacht, Regularity lemmas for graphs, in Fete of combinatorics and computer science, Bolyai Soc. Math. Stud., Vol. 20, 287–325, János Bolyai Math. Soc., Budapest, 2010.
- [104] V. Rödl and J. Skokan, Regularity lemma for uniform hypergraphs, *Random Structures Algorithms* **25** (2004), 1–42.
- [105] K. F. Roth, On certain sets of integers, *J. London Math. Soc.* **28** (1953), 104–109.
- [106] I. Z. Ruzsa and E. Szemerédi, Triple systems with no six points carrying three triangles, in Combinatorics (Keszthely, 1976), Vol. II, 939–945, Colloq. Math. Soc. János Bolyai, 18, North-Holland, Amsterdam-New York, 1978.
- [107] W. Samotij, Stability results for random discrete structures, *Random Structures Algorithms* **44** (2014), 269–289.
- [108] A. A. Sapozhenko, On the number of connected subsets with given cardinality of the boundary in bipartite graphs (in Russian), *Metody Diskret. Analiz.* **45** (1987), 42–70.
- [109] A. A. Sapozhenko, On the number of independent sets in extenders, *Discrete Math. Appl.* **11** (2001), 155–161.
- [110] A. A. Sapozhenko, Independent sets in quasi-regular graphs, *European J. Combin.* **27** (2006), 1206–1210.
- [111] D. Saxton and A. Thomason, Hypergraph containers, *preprint*.
- [112] M. Schacht, Extremal results for discrete random structures, *preprint*.
- [113] A. Scott, Szemerédi’s regularity lemma for matrices and sparse graphs, *Combin. Probab. Comput.* **20** (2011), 455–466.
- [114] M. Simonovits, A method for solving extremal problems in graph theory, stability problems, in Theory of Graphs (Proc. Colloq., Tihany, 1966), 279–319, Academic Press, New York, 1968.
- [115] J. Spencer, Ramsey’s theorem – a new lower bound, *J. Combin. Theory Ser. A* **18** (1975), 108–115.

- [116] B. Sudakov and V. H. Vu, Local resilience of graphs, *Random Structures Algorithms* **33** (2008), 409–433.
- [117] T. Szabó and V. H. Vu, Turán’s theorem in sparse random graphs, *Random Structures Algorithms* **23** (2003), 225–234.
- [118] E. Szemerédi, On sets of integers containing no  $k$  elements in arithmetic progression, *Acta Arith.* **27** (1975), 199–245.
- [119] E. Szemerédi, Regular partitions of graphs, in *Problèmes combinatoires et théorie des graphes* (Orsay 1976), Colloq. Internat. CNRS, 260, 399–401, CNRS, Paris, 1978.
- [120] T. Tao, A variant of the hypergraph removal lemma, *J. Combin. Theory Ser. A* **113** (2006), 1257–1280.
- [121] A. G. Thomason, Pseudorandom graphs, in *Random graphs ’85* (Poznań, 1985), North-Holland Math. Stud., Vol. 144, 307–331, North-Holland, Amsterdam, 1987.
- [122] A. G. Thomason, Random graphs, strongly regular graphs and pseudorandom graphs, in *Surveys in Combinatorics 1987*, London Math. Soc. Lecture Note Ser., Vol. 123, 173–195, Cambridge University Press, Cambridge, 1987.
- [123] Paul Turán, Egy gráfelméleti szélsőértékfeladatról, *Mat. Fiz. Lapok* **48** (1941), 436–452.
- [124] B. L. van der Waerden, Beweis einer Baudetschen Vermutung, *Nieuw. Arch. Wisk.* **15** (1927), 212–216.